

P1EsTHE1

1 UNITED STATES DISTRICT COURT
2 SOUTHERN DISTRICT OF NEW YORK
-----x

3 THE NEW YORK TIMES COMPANY,

4 Plaintiff,

5 v.

23 Civ. 11195 (SHS)

6 MICROSOFT CORPORATION, et al.,

7 Defendants.

8 -----x
9 DAILY NEWS, LP, et al.,

10 Plaintiffs,

11 v.

24 Civ. 03285 (SHS)

12 MICROSOFT CORPORATION, et al.,

13 Defendants.

14 -----x
15 THE CENTER FOR INVESTIGATIVE REPORTING, INC.,

16 Plaintiff,

17 v.

24 Civ. 04872 (SHS)

18 OPEN AI, INC., et al.,

19 Defendants.

20 New York, N.Y.
21 January 14, 2025
22 10:45 a.m.

23 Before:

24 HON. SIDNEY H. STEIN,

25 District Judge

SOUTHERN DISTRICT REPORTERS, P.C.
(212) 805-0300

P1EsTHE1

1 APPEARANCES

2 SUSMAN GODFREY LLP
3 Attorneys for Plaintiff New York Times
4 BY: IAN B. CROSBY

5 ROTHWELL FIGG ERNST & MANBECK PC
6 Attorney for Plaintiff New York Times and Daily News
7 BY: STEVEN LIEBERMAN
JENNIFER MAISEL

8 LOEVY & LOEVY
9 Attorney for Plaintiff CIR
BY: MATTHEW TOPIC

10 ORRICK, HERRINGTON & SUTCLIFFE LLP
11 Attorneys for Defendant Microsoft
12 BY: ANNETTE L. HURST
CHRISTOPHER CARIELLO

13 MORRISON & FOERSTER LLP
14 Attorneys for Defendant OpenAI, Inc.
15 BY: JOSEPH C. GRATZ
CAROLYN M. HOMER

16 LATHAM & WATKINS LLP
17 Attorneys for Defendant OpenAI, Inc.
18 BY: ANDREW GASS
ELANA NIGHTINGALE DAWSON

19 ALSO PRESENT:
20 KAREN A. CHESLEY, New York Times Counsel

P1EsTHE1

1 (Case called)

2 THE COURT: Counsel, make your appearances.

3 THE DEPUTY CLERK: Counsel, please state your names
4 for the record.

5 MR. CROSBY: Ian Crosby for The New York Times.

6 MS. CHESLEY: Karen Chesley for The New York Times.

7 THE COURT: Please speak into the microphone.

8 MS. CHESLEY: Karen Chesley for The New York Times.

9 MR. LIEBERMAN: Steven Lieberman for The New York
10 Times and Daily News plaintiffs.

11 MS. MAISEL: Jennifer Maisel for The New York Times.

12 MR. TOPIC: Matt Topic for Center for Investigative
13 Reporting.

14 THE COURT: All right.

15 Sorry. Just a moment.

16 MS. HURST: Good morning, your Honor.

17 Annette Hurst for Microsoft.

18 MR. CARIELLO: Chris Cariello for Microsoft.

19 MR. GRATZ: Joe Gratz for OpenAI.

20 MS. HOMER: Carolyn Homer for OpenAI.

21 MR. GASS: Andy Gass for OpenAI.

22 MS. DAWSON: Elana Nightingale Dawson for OpenAI.

23 THE COURT: Good morning, all of you. Please be
24 seated.

25 There are people listening on the listen-only line. I

P1EsTHE1

1 wish to remind those people that any recording of this
2 proceeding is illegal. Do not do it.

3 I have some thoughts on how we should proceed this
4 morning. I also wanted to see if counsel had talked to each
5 other about how to proceed this morning. My own thinking is if
6 there is a unified view by counsel here, I'll certainly listen
7 to it. My only thinking is that it makes sense to go claim by
8 claim rather than case by case. I'll tell you the order of the
9 claims that I have. That's one point.

10 The second point is, it seems to me appropriate, if
11 the parties wish, to make presentations on how the technology
12 works. I think I have a basic understanding of it. Each time
13 I go through it I have a better understanding of it. That will
14 be up to the parties as to whether they want to do that. And
15 it seems to me that that probably should be at the beginning.

16 Can somebody talk to me. Have you all talked about a
17 presentation today or not?

18 You're first in line.

19 MR. CROSBY: Yes.

20 THE COURT: By the way, when you speak to me, please
21 say your name and who you represent. I haven't met any of you
22 before, I don't believe, and there are a number of parties
23 here.

24 MR. CROSBY: Yes.

25 Ian Crosby for The New York Times, your Honor.

P1EsTHE1

1 We have not conferred about how to proceed, but it was
2 our intention and our hope that the court would proceed issue
3 by issue, and we are prepared to do so on the plaintiffs' side.

4 THE COURT: Is there anyone who disagrees with that?

5 All right. What about a presentation on how these
6 things work?

7 Does anyone think that's necessary, unnecessary, when
8 should it come?

9 MR. HURST: Your Honor, Annette Hurst for Microsoft.

10 The defendants think a brief discussion tutorial about
11 the technology would be appropriate.

12 THE COURT: Yes.

13 Sir.

14 MR. CROSBY: We believe so, as well. Ms. Maisel is
15 prepared to do so, and we would like to go first as the
16 plaintiffs.

17 THE COURT: Fine. Then I'll let the defendants go
18 first when we turn to the motions because they are the ones who
19 have made the motions.

20 All right. We have a way of proceeding.

21 All right. Ms. Maisel, you're on behalf of The Times
22 and the Daily News, correct?

23 MS. MAISEL: That's correct, your Honor.

24 THE COURT: Why don't you talk to me.

25 You can speak from there, you can use the podium,

P1EsTHE1

1 whichever you prefer, but I do ask that you speak loudly,
2 slowly, and clearly into the microphone.

3 MS. MAISEL: OK, your Honor. Well --

4 THE COURT: We're having trouble making sure
5 everyone on the line can hear, which is why the IT people are
6 coming up.

7 MS. MAISEL: All right. For brief technical
8 background and as it relates to the claims in the case. A part
9 of this case concerns largely which --

10 THE COURT: I'm sorry. Can you speak more loudly.

11 MS. MAISEL: OK. At the heart of this case are large
12 language models. These are artificial intelligence models that
13 are designed to take as input text content and output text
14 content. And as it relates to the claims of the case, there
15 are really four different areas of these large language models
16 and the technology surrounds them that are at issue.

17 At the first stage, whenever we think about artificial
18 intelligence, we have to follow the data, similar to criminal
19 cases where you follow the money. And if you look at the data,
20 at the first instance, we as the news plaintiffs, content is
21 published online, it's published on the web. And the
22 defendants' crawlers, their box, these automated tools are used
23 to collect that content, to scrape that content.

24 And scraping involves accessing web pages, sometimes
25 behind the pain wall, and copying that content, cleaning it up.

P1EsTHE1

1 In some cases that includes stripping content management
2 information, which is at issue in the case, and storing copies
3 of those content in training datasets or training corpora.

4 This first stage of the process, this curation of
5 training datasets, we call this ingestion. So the process of
6 collecting data and creating training datasets.

7 The second phase --

8 THE COURT: When you say training datasets, I've seen
9 in all of the materials I've been reading, that the information
10 is stored as packets.

11 Is that little bits of information, is that how it's
12 stored?

13 MS. MAISEL: So, in the very -- when we refer to
14 training datasets or training corpora, we are referring to the
15 actual text content, so it's perceptible to humans. You can
16 open a file on the training dataset and read its contents. So,
17 technically, it is stored as bits, just as any other data is
18 stored on a computer, but it's perceptible to a human.

19 But the second phase, which I think your Honor might
20 be thinking about, this digestion phase, this is the processing
21 of those training datasets, how those training datasets and
22 each piece of content within those training datasets are fed
23 through the models, through algorithms, as part of the training
24 process. And we call it digestion because as part of the
25 training phase, one phenomena we observed -- and this is

P1EsTHE1

1 attached as Exhibit J to both the New York Times and the Daily
2 News complaint -- because of the technology, how it works, and
3 the training process. For one reason or another, content is
4 encoded or stored within the models themselves as part of this
5 training process. That is through the different parameters of
6 the models.

7 And just to take a step back, these models involve
8 lots and lots of parameters. Each parameter or weight is a
9 numerical number, and those numerical numbers are updated as
10 part of the training process. So the model, it's a set of
11 numbers, set of model parameters, and training data is fed
12 through different algorithms in order to adjust those
13 parameters.

14 As part of this process --

15 THE COURT: Can you give me a real-world example of
16 that.

17 MS. MAISEL: Of the parameters?

18 THE COURT: Of the process you're talking about. I
19 understand that you have these bots that scrape -- not you, but
20 the defendants -- that scrape a great deal of information from
21 the web and wherever they can find it. In fact, trillions of
22 words, apparently. So all of this is fit into the large
23 language models, the LLMs. They then do something to it.

24 What is it that they do?

25 And I guess we can talk separately about whether or

P1EsTHE1

1 not they strip the CMI, but it's probably for later in this
2 discussion.

3 But go ahead.

4 MS. MAISEL: So, your Honor may be referring to the
5 concept of tokenization. So each piece of training data --
6 this is a piece of text in the case of large language models --
7 it is tokenized, meaning it is broken up into small fragments,
8 a token. We could think of it as a word.

9 THE COURT: That's what I was thinking of before.

10 Go ahead.

11 MS. MAISEL: Yes. It could be a word or a fragment of
12 a word. Just a handful of tokens, each token corresponding to
13 a character.

14 Each piece of training data is tokenized. It's broken
15 down into an embedding or a vector, a mathematical
16 representation of that, of those tokens, and it's fed to the
17 model. And the training process is designed so as these words
18 are broken up into tokens, the relationships between the words,
19 what words come, what tokens come before other words, what
20 tokens come after, how frequently tokens come up, all of this
21 information is fed through the parameters of the model as part
22 of training.

23 And the ultimate goal of this training process -- and
24 I'm referring to pretraining here, there a couple of different
25 phases.

P1EsTHE1

1 THE COURT: Did you say pretraining?

2 MS. MAISEL: Pretraining. There is post training,
3 which we can talk about.

4 But as a part of pretraining, the training corpus,
5 each document is broken up in a token, fed through the model.
6 And the ultimate goal is to train the model to predict the next
7 word to come after any given word. The next token that comes
8 after a specific token. That is the goal, as we understand of
9 it, of the training process.

10 This is what we call, again, going back to our
11 analogy, the digestion phase. You have ingestion, training
12 datasets, feeding it to the model. Digestion. All this
13 training. And in some instances, training occurs multiple
14 times with respect to any given piece of content.

15 As detailed in the complaint, we describe how, in some
16 of the defendants, particularly OpenAI's research papers, they
17 sometimes go over the same piece of training content multiple
18 times. Starting content can be weighed more heavily during
19 training, particularly high news -- high-value content, high-
20 accuracy content.

21 So, during this pretraining process, you know, based
22 on the allegations in our complaint and discovery so far,
23 higher quality content is weighted more heavily during
24 training. Again, the ultimate goal is we want the language
25 models to build and accurately and competently redirect the

P1EsTHE1

1 next word in a sequence. So that is digestion.

2 We then get regurgitation. As detailed in both
3 complaints, the Daily News and The New York Times complaint,
4 and particularly in Exhibit J, there is this phenomena where,
5 in response to the right prompt, so the language model, the
6 language model will spit out copies of training data that it
7 saw.

8 THE COURT: Copies of training data that it?

9 MS. MAISEL: That it saw during training.

10 THE COURT: Go ahead.

11 MS. MAISEL: In response to Exhibit J, there is a
12 small input text of the beginning of an article and the model
13 will complete that article without referencing.

14 THE COURT: What is the first paragraph, what's the
15 next paragraph, that's what you're talking about?

16 MS. MAISEL: Yes, exactly.

17 THE COURT: OK.

18 MS. MAISEL: That's due to a phenomena called
19 memorialization. And a phenomena of outputting that training
20 data, we call regurgitation. So then that's the training
21 process in a nutshell.

22 There is post training, which I don't think we need to
23 get into for purposes of the claims before us right now. But
24 there is another set of technology called retrieval augmented
25 generation. This is a technology that sits outside of the

P1EsTHE1

1 training process. As detailed in the complaint, retrieval
2 augmented generation involves roughly a five-step process.

3 It's in paragraph 78 of the Daily News complaint.

4 THE COURT: Just a moment. Let me turn to it, if
5 you're quoting it. I'm familiar with the law, but let's turn
6 to 78.

7 MS. MAISEL: It's also 81 of The New York Times
8 complaint.

9 THE COURT: All right. Go ahead.

10 MS. MAISEL: So I detailed the three steps of training
11 and one issue of training can take a lot of time. It can take
12 a lot of compute power. It can take several days or weeks or
13 months to train a model, especially the ones as large at issue
14 in the case.

15 As a result, we sometimes see something called a
16 knowledge cutoff date. So, at some point the training data
17 stops in time and the model, as it's trained, doesn't know
18 anything beyond that knowledge cutoff date.

19 So retrieval augmented generation, RAG, or as we
20 detail in the complaint, we call it synthetic search. This is
21 a process where external data sources may be referenced or
22 retrieved in order to augment the language model's knowledge
23 base. And as detailed in paragraph 78 of the Daily News
24 complaint, there are five steps here, or four steps.

25 First step is a user puts in a prompt. A prompt could

P1EsTHE1

1 be, tell me about the news today, or tell me about a specific
2 article, give me the first paragraph of an article. That's a
3 prompt.

4 The defendants' products, and here I'm talking about
5 OpenAI's ChatGPT with the browse plug-in, or some other custom
6 GPTs. Or in the case of Microsoft, their Copilot chat program
7 or Bing chat program. These products are coupled with
8 Microsoft's Bing search engine. What happens is the prompt
9 that is sent to one of these products, one of the defendants'
10 products is translated into queries, queries that are then sent
11 to the search index. Just like research, a user would go to
12 the Bing search feature or to Google or any other search
13 engine. These queries are sent to the search index and certain
14 responses, web pages, are returned in response to these
15 prompts. And those responses include copies of content of web
16 pages.

17 So, in the case of a news article, what might be
18 returned from the search index is a copy of an article, copy of
19 a website that contains that article. And that content, the
20 copy of the content or the news article that's returned from
21 the search index is then provided together with the prompt to a
22 large language model, to the GPT model.

23 So just to drive this home, as an example, I, as a
24 user, might put in a prompt, tell me about the first paragraph
25 of a specific article. The products will translate that to a

P1EsTHE1

query that can be interpreted by a search index. The search index returns a response that includes something that hits on that prompt so the article and the contents of that article. And my question, you know, what is the first paragraph, along with the content of that article, are provided to the LLM. Say, give me the first paragraph of this article. Here is the article.

The LLM, in turn, returns a response. And that response here is, it might return the verbatim first paragraph, it might return several paragraphs or, as detailed in the complaint, sometimes the entire article is returned to the user in response to that prompt. But the difference here with retrieval augmented generation is because that content, that very fresh content was not seen during training because of that knowledge cutoff, we're talking about something that was not used during training, but is still used in connection with these large language models, coupled with the Bing index, in order to provide a very detailed, satisfied summaries of verbatim quotes or sometimes the entire article back to users.

THE COURT: I'm not quite sure what this last point is. I understand how you put a prompt in. Tell me what the first paragraph of the article in The New York Times today about the Hegseth confirmation hearing. I then will get, as the end user, I then will get back, presumably, the first paragraph of an article in The New York Times.

P1EsTHE1

1 But what was the last point you made? I didn't
2 understand that.

3 MS. MAISEL: In some instances, it's not the first
4 paragraph that's returned.

5 THE COURT: This is what I didn't understand. The
6 difference here with retrieval augmented generation is because
7 that content, the very fresh content that was not seen during
8 training, that's what I don't understand.

9 MS. MAISEL: Yes. So take your example with the
10 article posted today. It's such fresh content that the
11 defendants could not have possibly have had time to train their
12 model on that content to incorporate that content as part of
13 the training process. They have to scrape the content, collect
14 it, these algorithms, to train the model. It's a lot of
15 compute power required.

16 So, as a result, there is some cutoff date in time for
17 what's part of the enclosed knowledge of that model after it's
18 trained.

19 THE COURT: My confusion is, I thought I read that the
20 models are always learning the more queries are made. They
21 get, I don't want to say, smarter, but they understand what, in
22 better ways, what words follow what words.

23 Am I wrong about that?

24 MS. MAISEL: I don't think you're wrong, your Honor.
25 I would say that's still an active part of discovery here, how

P1EsTHE1

1 user queries, user feedback are used as part of the training
2 process.

3 THE COURT: OK.

4 MS. MAISEL: That might be a good question for the
5 defendants.

6 THE COURT: OK. Go ahead. I didn't mean to stop you.

7 MS. MAISEL: Yes.

8 THE COURT: I'll be very interested in the view of the
9 defendants on how these things work is different.

10 But go ahead.

11 MS. MAISEL: I mention that sometimes -- so, take this
12 example, give me the first paragraph of an article. Sometimes
13 the first paragraph is returned in its true and faithful form.
14 Sometimes a summary is returned. Sometimes something
15 completely made-up and fabricated is returned.

16 THE COURT: That's your hallucination issue. The
17 papers say the parties are working, the defendants, are working
18 to eliminate.

19 MS. MAISEL: As they allege, yes.

20 As detailed in the complaint and for purposes of the
21 copyright infringement claim, through this retrieval augmented
22 generation process, copies of the copyrighted works are created
23 throughout this process, created as part of the search index,
24 they are created as part of the response from the search index,
25 and they are created as part of the prompt to the large

P1EsTHE1

1 language model. And sometimes they are copied further as
2 output to end users using these products.

3 That is retrieval augmented generation, which it's
4 coupling search index technology, the Bing search index from
5 Microsoft, with the large language models at issue.

6 Does your Honor have any further questions?

7 THE COURT: No. I appreciate your presentation.

8 Thank you.

9 MS. MAISEL: You're welcome.

10 THE COURT: Who would like to speak on behalf of the
11 defendants?

12 Am I going to hear somebody both from Microsoft and
13 from OpenAI, or did you want to make a joint presentation?

14 MR. GRATZ: You are going to hear from both of us,
15 your Honor.

16 This is Joe Gratz for OpenAI. I'm going to cover the
17 training-related issues, the creation of the large language
18 model, and Ms. Hurst will cover the Next Generation search
19 engine technology that Ms. Maisel also talked about, the thing
20 that happens after training, the retrieval of generation
21 technology will be addressed by Ms. Hurst.

22 THE COURT: OK.

23 MR. GRATZ: So, your Honor, we agree conceptually with
24 many, though not all, of the things Maisel, Ms. Maisel,
25 described.

P1EsTHE1

1 What I do want to start by talking about is the
2 separation between two parts of this process. The part of
3 training the model, that is, the creation of this large
4 language model through the process of training by having it
5 effectively read or train on trillions of examples of language,
6 to learn the relationships between words and concepts by
7 reading, sort of, everything. And then after the model is
8 created, the process of providing prompts, getting outputs, and
9 things that you can do in that process, after the model has
10 been created, while you're using the model, that the retrieve
11 augmented generation, the Next Generation search engine
12 functionality, that happens after the training has happened.

13 So, to return to your Honor's question about are these
14 models, sort of, effectively training themselves all the time?

15 Are they just getting smarter, everything we say to
16 them, piece by piece?

17 The answer is the training process is, sort of, a
18 process that occurs once to create a particular version of a
19 model, and there is a set of training data on which the model
20 is trained, those trillions of examples. As one is interacting
21 with the model, the model is not sort of changing itself or
22 retraining itself live. It's having --

23 THE COURT: It's following the algorithm to spit out
24 this word, then that word, and then the next word.

25 MR. GRATZ: That's right. It is predicting the next

P1EsTHE1

1 word in the conversation at that time based on its training.
2 So that separates out those two situations.

3 THE COURT: So is the universe of data that's used for
4 GPT-2 different and smaller than the universe of data used to
5 train GPT-4?

6 MR. GRATZ: Yes.

7 THE COURT: OK. Go ahead.

8 MR. GRATZ: That's right.

9 What these large language models are made for is to be
10 able to generalize, to understand relationships between
11 concepts, to understand language, to understand the
12 relationships between words, to understand facts. And the only
13 way to do that is to provide many, many, many trillions of
14 examples of uses of language and statements about the world in
15 the form of this massive amount of training data.

16 Training consists of the model learning from those
17 examples to be able to use language itself. The point of this
18 is to create a general purpose language-using tool that can
19 be -- can respond to any prompt or question with some kind of
20 language response, just as another language user, like a person
21 can respond to any prompt based on what they have been exposed
22 to in their life.

23 So you can say, hello, and it will predict that the
24 next thing somebody else might say is, how are you today?

25 But it goes beyond that. You can say, What's the

P1EsTHE1

1 capital of France? And it can say, The capital of France is
2 Paris. Right, because that is a fact that it has learned from
3 the world. You can do lots of things other than retrieving
4 facts or data.

5 You can say, Here is a letter to my landlord. Can you
6 make it sound less angry. And it will do that because it has
7 before it what you have written and it can transform that into
8 a less angry version of that letter to your landlord.

9 You can say, Write me a poem about the architecture of
10 500 Pearl Street, and it knows things about 500 Pearl Street
11 because it's read them, knows things about poems, and knows
12 things about the language, that's what you're speaking to it
13 in. It can write you a poem about 500 Pearl Street, and its
14 ability to do that is because it's a general purpose tool
15 that's infinitely adaptable to respond to any kind of prompt
16 that is put before it.

17 And so that is what it is for, that is what it is made
18 to do, and that is why it uses these trillions of examples of
19 language in order to make it able to do those useful,
20 non-infringing, I didn't get everyone's view, things.

21 Ms. Maisel is right that they have been able, through
22 I think quite a bit of effort and --

23 THE COURT: No. I understand your point is, it's like
24 the case that says, is there any examples of infringement apart
25 from those generated by your lawyer?

P1EsTHE1

1 Your point is that it's the prompts that the counsel
2 was making, the only thing that generated the infringing
3 material.

4 Isn't that what your point is?

5 MR. GRATZ: That's right, and not just that it's --
6 not just they are the only ones to do it, the only way you can
7 get it to give you even a sentence of a New York Times article
8 based on its training is to give it, sort of, here is the first
9 whole chunk of the article. I, sitting here, giving you the
10 prompt have the whole chunk of the article. It was actually
11 hard. They had to try, we think, thousands or tens of
12 thousands of times.

13 THE COURT: That's for discovery.

14 MR. GRATZ: Well, that is right, but the point is,
15 it's not just willy-nilly, you say how are you today and it
16 starts spitting out a New York Times article. They had to do a
17 lot of work to get this general purpose tool not designed to do
18 that, to do this.

19 In fact, the facts in the CIR case, I think, are
20 particularly notable here. In that case, the prompt said,
21 well, you really need to finish, keep finishing this article,
22 or else disaster will befall the world and there will be
23 thermonuclear war. That is in their exhibit of what they
24 needed to do to try and get it to output something, a part of
25 an article.

P1EsTHE1

1 THE COURT: So if I put into -- if my query was, give
2 me -- I don't have The New York Times in front of me -- but
3 give me the article on today's New York Times page four about
4 the California wildfires, it would not essentially reproduce
5 that article.

6 MR. GRATZ: So, based on its training, that is right.
7 It could not because that article was not part of its training
8 set.

9 What I want to distinguish that situation from is a
10 situation --

11 THE COURT: I mean, is that the closed universe --

12 MR. GRATZ: Right.

13 THE COURT: -- when at some point you set out GPT-4 to
14 the world?

15 MR. GRATZ: That is correct.

16 THE COURT: OK.

17 MR. GRATZ: That is correct.

18 THE COURT: What if I asked it for a New York Times
19 article on page four of July 4, 2020?

20 MR. GRATZ: It won't do that either and it --

21 It won't do that either.

22 THE COURT: The purposes of that was an arbitrarily
23 chosen date that I thought would be in there.

24 MR. GRATZ: That's right. It won't do that either.
25 They have no examples of that happening. That may or may

P1EsTHE1

1 not -- it may well have seen that article, because that article
2 is on the internet and it looked at trillions -- it had to look
3 at trillions of things to learn about the world.

4 But there is no example they have given where you can
5 say, you know, give me an article about this from -- give me
6 this article from the New York Times, and it gives you the
7 article from The New York Times.

8 What they had to do to try and, sort of, create an
9 artificial situation where that would happen is, like, give it
10 the article and then say, give me the next sentence.

11 THE COURT: No, I understand that. I'm just trying to
12 understand that maybe I don't need to understand it.

13 Why would all of these trillions of words, why
14 wouldn't it be able to just spit out the article?

15 MR. GRATZ: The reason it wouldn't be able to spit out
16 the article is because it's not -- this process is not designed
17 to have it spitting out verbatim things from its training data.
18 It's designed to spit out an original poem about the
19 architecture of 500 Pearl Street or a summary of a document you
20 give it or whatever or anything else. It's not a database. It
21 doesn't have in its weights all of its training data. It has,
22 the training data has the inferences and information that is,
23 sort of, gleaned from having been exposed to those trillions of
24 tokens.

25 THE COURT: Again, just -- I'm sorry I'm so basic

P1EsTHE1

1 here.

2 It does, presumably it does have that New York Times
3 article in its database, correct?

4 MR. GRATZ: No, it does not. There is not a
5 database -- and I think Ms. Maisel's description of this was
6 technically accurate -- which is it is a set of weights and
7 that is numbers, and those numbers are updated and affected by
8 the training data. And those numbers are a way of predicting,
9 sort of, continuing a conversation.

10 Those numbers, when you run them in the inference
11 process, allow it to predict the next information, the
12 information that is about the content or gleaned from reading
13 the content. It's not a literal representation of any of the
14 content.

15 And part of -- I mean, and why is that?

16 It's because this isn't an information -- this isn't
17 like a document retrieval system. There are lots of very good
18 document retrieval systems out there. This is a language
19 model, and the fact that the language model sometimes picks the
20 next word from something it has seen is true of any language
21 user.

22 If I say to you, you know, yesterday all my troubles
23 seemed so, you know, we would all think to ourselves, far away,
24 because we have been exposed to that text so many times. That
25 doesn't mean you have a copy of that song somewhere in your

P1EsTHE1

1 brain. And to the same, by the same token, it doesn't mean
2 that the weights of the model are a database of information.
3 It's just those -- the inferential information in there leads
4 it to sometimes, based on patterns it has seen, pick the next
5 sequence of tokens, particularly in the artificial situation
6 that they have created where they say, here is, you know, here
7 is a Pulitzer-winning article that is all over the internet.
8 Can you guess the next word? And it can guess the next word.
9 Not because that is what it is supposed to do, that is just
10 what it happens to do because it is out there to pick the next
11 token.

12 Now, to be clear, because that is not what we want it
13 to do, this is not in the pleadings, but there are lots of ways
14 that we work to keep it from doing that and because that is not
15 its intended behavior. It is not intended to be a database
16 where you can come to it with the first paragraph of The New
17 York Times article and get most of the next paragraph from The
18 New York Times article, or in CIR's case, first paragraph of a
19 Mother Jones article and get the next ten words of that Mother
20 Jones article. That is not what it is for, that is not what it
21 is designed to do, and that is not what it does, unless you're
22 sitting at that table.

23 Those are the, sort of, key points that I wanted to
24 make about the training process. Unless there is anything else
25 your Honor wants to discuss on that, I will turn it to

P1EsTHE1

1 Ms. Hurst with respect to the search engine.

2 THE COURT: Thank you.

3 MR. HURST: Good morning. Annette Hurst from
4 Microsoft.

5 From Mr. Gratz' description, you can see that this is
6 a general purpose technology. Large language model embodies
7 the statistical relationships of words to one another across
8 billions of context.

9 So, your Honor, the model understands that a
10 strawberry, when ripe, is sweet. It understands Strawberry
11 Fields Forever, all the context in which the word strawberry
12 might occur. Those set of probabilities about the occurrence
13 of the word strawberry, a variety of context is what is
14 embodied in these models.

15 Now, your Honor, the model, the large language model
16 which has that understanding of language can be adapted once
17 it's trained to a wide variety of purposes, everything from
18 curing cancer to national security. And that is why in these
19 complaints we see in The New York Times complaint, at paragraph
20 61 to 63 and 81, in the New York Daily News complaints,
21 paragraphs 58 to 60, a description of the substantial
22 commercial non-infringing uses of this technology.

23 In other words, the plaintiffs, in their own words,
24 have alleged that this technology is capable of being
25 commercialized to the tune of billions of dollars without

P1EsTHE1

1 regard to any capability for infringement. When they are
2 complaining about Bing chat, what they are alleging is that one
3 implementation of this general purpose technology is a Next
4 Generation search engine. Now, we know in this Circuit from
5 the *Authors Guild* cases that the search engine is fair use. So
6 they are coming into this with a heavy burden of showing that a
7 Next Generation search engine doesn't fall within that
8 category.

9 But what is so great about Bing chat and these Next
10 Generation search engines?

11 What is great about them is that they are actually
12 designed to answer your question, your Honor. So it's not just
13 giving you a series of links to things you can go look for
14 yourself and maybe find the information. But one of the great
15 uses of a large language model is that it's able to interpret
16 and summarize and derive facts from underlying data and then
17 deliver those facts to you.

18 Now, the delivery of facts is a historically protected
19 pro-copyright purpose for transformative technology, and that
20 is exactly what this Next Generation search engine is doing.

21 Now, does it generate a copy in the context of doing
22 that? Sure, it does. As Ms. Maisel described, it takes the
23 traditional Bing search index, which is an index of articles
24 that is used to return, in a fair use process, every single day
25 links to a whole wide variety of content on the internet. All

P1EsTHE1

1 the publicly accessible content on the internet.

2 And the large language model then summarized, finds
3 and summarizes the pertinent articles and delivers an answer to
4 the court. You know, for example, my college-age daughter was
5 having tailbone pain. And so I asked Copilot, what are some of
6 the causes of tailbone pain? And it went out and looked at the
7 Mayo clinic and Med Page news and a whole wide variety of
8 sources, and it came back to me and it gave me a list which was
9 really useful. I can help her, talk to her doctor about what
10 might be the problem.

11 And so this Next Generation search engine, which is an
12 implementation of the LLM technology, helps people answer
13 questions in ways that are more concise, more factual, and for
14 more pertinent than the traditional search engines.

15 THE COURT: Help me a little on that. What you
16 posited sounded like, to me, questions that are put into search
17 engines every day. So what is new about this large language
18 model?

19 Bing presumably has always been able to print out
20 articles on tailbone pain. What are you telling me is new?

21 MR. HURST: Your Honor, what is new is when you, in a
22 traditional search bar, you put in a query, you get a series of
23 links. In the Next Generation search, Copilot chat, you ask a
24 question in a natural language format and it gives you an
25 answer in a natural language format with cites links to those

P1EsTHE1

1 sources.

2 So, in essence, the large language model is
3 summarizing the information that it's found for you and
4 delivering you the factual content of that information, and
5 then it's giving you citations and the sources for that
6 information.

7 THE COURT: So is the traditional way simply going to
8 give me a list of sources, list of articles on tailbone pain,
9 but the new way, the Next Generation is going to give me a
10 plain language answer that is supported with essentially the
11 same links?

12 MR. HURST: That's right, your Honor.

13 And what the large language model is capable of doing
14 that the traditional search models, which by the way were also
15 predicative algorithms and models, is it's better at
16 interpreting your query because it understands language better.
17 It's better at knowing exactly what you're asking. And it's
18 also capable of putting all of the data together and giving you
19 an answer, rather than just giving you a series of sources.

20 THE COURT: OK. All right. That's helpful.

21 MR. HURST: I think that's all I would emphasize about
22 the generative search technology, your Honor, unless the court
23 has any questions.

24 THE COURT: No. Thank you. I think this was helpful.
25 I appreciate it.

P1EsTHE1

1 So why don't we turn now to...

2 MR. CROSBY: Your Honor, there is a couple of points
3 that I think are still in dispute that ought to be clarified in
4 response to that presentation.

5 THE COURT: You mean in dispute about how these things
6 work?

7 MR. CROSBY: Yes, fundamentally about how these things
8 work and how that relates to the well-settled law. I just
9 wanted to pick up on --

10 THE COURT: You're Mr. Crosby for The New York Times.

11 MR. CROSBY: What Mr. Gratz said, that these models
12 learn facts, that what they are doing is they are taking all of
13 this information in these trillions of tokens derived from all
14 of the text that is out there on the internet, and they are
15 abstracting facts and concepts. This is a fundamental point of
16 dispute in this lawsuit.

17 What we contend is that is exactly not what large
18 language models do. And there are any number of examples that
19 are out there that demonstrate why this is not the case. They
20 can be read every chess book that's ever been written, but they
21 can't play chess. The example that was given about Paris is
22 the capital of France. There is a famous paper called *The*
23 *Reversal Curse*, where if you train a model -- this is an
24 experiment -- if you train a model only using the form of the
25 expression Paris is the capital of France, and if you ask it,

P1EsTHE1

1 you know, the capital of France is Paris, and you ask what is
2 the capital of France, it will say Paris. If you ask it, what
3 is Paris? It won't know what the capital of France is. It
4 won't say the capital of France. Because it only absorbs the
5 expression of the fact, it never learns the underlying fact.

6 These are just statistical models of how language is
7 presented of the expression, the expression which is what
8 copyright law protects, and they have not bootstrapped their
9 way into actually understanding what the text means. And that
10 is why this problem of hallucination exists and why it has been
11 so intractable and why it's not been solved.

12 Because these models, they don't have our ability to
13 read something and actually know what the underlying
14 information is. They can just sort of go off stringing out one
15 word after another, with no notion that I saw something that
16 contradicts what I'm saying right now. They don't have that
17 ability.

18 The other thing that I really --

19 THE COURT: Now, what does that do for you?

20 MR. CROSBY: Well, I think what the argument is,
21 is that, well, what we're -- even though we are copying the
22 expression, what we are really after is the non-copyrightable
23 underlying facts and ideas. And just because of the way the
24 computers process information, the only -- you and I can look
25 at a page and we can learn something in our mind and that is

P1EsTHE1

1 not considered to be a copy, under the definition in 17 U.S.C.
2 101, because you can't retrieve it or perceive it. I can't
3 read your mind, right?

4 And, so they want to say, this is analogous to what a
5 human does and, therefore, it should be no more considered to
6 be copyright infringement than when you and I read something
7 and learn something because the copying is sort of incidental
8 to this fair use or non-copyright protected purpose of learning
9 facts.

10 But that is actually not, we believe, the case. It is
11 the expression and the expression of the word, the facts, and
12 not the facts that are embodied in these parameters of the
13 model and that is what is recalled when the models are prompted
14 to elicit text at one time.

15 The other thing that I wanted to emphasize, I was
16 surprised to hear Ms. Hurst say this, that the term that
17 Ms. Hurst was using was that these are Next Generation search
18 engines. Our contention is actually that they are not Next
19 Generation search engines, but that they are answer engines,
20 and that is, in fact, what Ms. Hurst said.

21 What the difference is between a previous search
22 engine and the Next Generation search engines, i.e. the answer
23 engines, is that instead of giving of you ten blue links back
24 to the original source where you can go and you can read it and
25 the publisher can potentially monetize that through advertising

P1EsTHE1

1 or through a pay wall and a subscription, which is the basic
2 bargain of the internet, intention for creative works, we are
3 just going to give you the answer.

4 OK. And that is the key fact in the fair use cases
5 and the search engine cases. The reason that copying the
6 entire internet to create a search index was deemed to be fair
7 use in the first place is because it is not substitutional.
8 Because the purpose is to allow you to find the original work.

9 And so here, you know, the argument, well, because we
10 are learning facts from the works and then giving you the
11 answer, that this is somehow footnotes with those cases. This
12 is the key distinction in this case. And that is what our
13 case, with respect particularly to these generated outputs is
14 about, because there are any number of studies about what the
15 impact of generative search is going to be on the market for
16 original creative work.

17 And the predictions are dire that, you know, that 30
18 to 50 percent of the traffic that is generated from search
19 engines back to original publishers is going to be diverted,
20 notwithstanding the fact that they put footnotes on these
21 answers. If anything, because people are aware of the problem
22 of hallucination, it may very well be that putting footnotes on
23 the answers that come out of the answer engine actually creates
24 more likelihood that people will not go back to those original
25 sources. It gives them confidence in the answer.

P1EsTHE1

1 So these are, obviously, we are far beyond the issues
2 that have been raised in the motions that the court is hearing
3 today, but the defendants used a large portion of their
4 briefing, sort of, trying to make their jury argument on the
5 fair use issue that is the heart of this case.

6 And so those particular points are going to be hotly
7 contested. There is going to be expert testimony and lots of
8 evidence about the potential for substitution of these answer
9 engines and what it is that these models actually do when they
10 ingest all of this valuable commercial copyrighted content to
11 train themselves to output text that can compete with that very
12 same content on the internet in another context.

13 Thank you, your Honor.

14 THE COURT: We are way ahead of ourselves.

15 Sir, I want to get to the issues.

16 MR. GRATZ: I do, too, your Honor. I want to state
17 this as briefly as I can in response to Mr. Crosby's comments
18 on the training process and what the model is doing.

19 It is not just understanding relationships between
20 words and sequences of words. It is understanding ideas. We
21 can see that by just using it. We can see that it knows what
22 Paris is. I just asked it. It's told me. It knows and it
23 understands the concept of Paris, the concept of New York City,
24 and how those things relate to one another. If you ask it,
25 compare Paris to New York City, it will write a new text,

P1EsTHE1

1 understanding those concepts, and talking about how they relate
2 to one another.

3 It is not an engine for reproducing expression. It is
4 an engine for producing new expression, its own expression,
5 containing ideas, perhaps, and facts that are in other
6 preexisting works, but creating something itself that is new.

7 On Mr. Crosby's view of the world, it would not be
8 able to do these things, and that is why Mr. Crosby's view is
9 wrong.

10 THE COURT: All right. Let's go to the issues here.

11 This is what I want to do. Then if I've left any live
12 issues out, you can tell me.

13 First, we'll hear on the first issue will be the claim
14 for direct copyright infringement in the statute, basically
15 limited to the statute of limitations argument. That's Count
16 One for The Times and for Daily News.

17 Secondly, contributory copyright infringement. That
18 is Count Four for The Times and for the Daily News, and Three
19 for the Center for Investigative Reporting.

20 Then the DMCA claim, that is the Digital Millennial
21 Copyright Act claim, which is Count Five for The Times and
22 Daily News, and a variety of news for Center of Investigative
23 Reporting.

24 Then unfair -- I'm sorry -- common law unfair
25 competition by misappropriation, which is Count Six for The

P1EsTHE1

1 Times and the Daily News.

2 Then the enrichment issue, that is direct copyright
3 infringement regarding these infringements. That's just the
4 Center for Investigative Reporting, Count Three.

5 Then New York State trademark delusion, trademark
6 delusion. That's Count Eight for the Daily News.

7 And last on my list is federal trademark delusion,
8 count Seven for the Daily News.

9 All right. Let's start with direct copyright
10 infringement in terms of the statute of limitations. Let me
11 hear first from, on the motions, I'll hear first from the
12 defendants, whichever defendant has raised it.

13 I would really like to keep down the amount of
14 duplication, so if one of the defendants says basically all
15 that is necessary, then let's not have duplication. It seems
16 to me that this is quintessentially an issue of fact that I
17 can't say that, as a matter of law, when the parties became
18 aware of the infringement, especially if I'm going to use the
19 test of, there has to be no doubt as to when the parties became
20 aware of it.

21 So the point here is, I think we're dealing with a
22 question of fact. What defendant wants to tell me I'm wrong?

23 MS. HURST: Your Honor, Andy Gass from Latham &
24 Watkins on behalf of OpenAI.

25 We can keep this one pretty short. We've got a long

P1EsTHE1

list of topics to get through here. We don't need to belabor the statute of limitations one. I think the legal question presented is simply whether The New York Times and the other plaintiffs were on what the law might call inquiry notice of the alleged infringement.

And the point is simply that the following facts are undisputed:

Fact one, as early as 2019 and 2020, more than three years ago, OpenAI published papers saying that among the content in the training set was New York Times articles.

Fact two, in July of 2020 --

THE COURT: Wait just a moment. Let me think about that.

Go ahead.

MS. HURST: Fact two, this was not or should not have been unknown to The New York Times and the other plaintiffs, because in July of 2020, also more than three years before the complaint was filed, The New York Times itself published an article about OpenAI's revolutionary generative AI technology.

If that were not enough, fact three, in December of 2020, both The New York Times and The Chicago Tribune, along with the Daily News plaintiffs, published another article, a lengthy one, discussing not just what the technology is, how it works, but also explicitly referencing the fact that trillions and trillions of words were used to generate the large language

P1EsTHE1

1 model at the heart of the technology.

2 Now, your Honor, if you don't believe that a
3 reasonable party at that point who himself published an article
4 describing how the technology works would not have an
5 obligation to investigate, I don't know when you would ever
6 find inquiry notice.

7 THE COURT: Is the test inquiry notice, or is the
8 test, the test that I have from the Southern District case
9 *PK Music Performance*, where the evidence has to eliminate all
10 doubt?

11 MS. HURST: Your Honor, I don't believe that is the
12 standard.

13 I think if you look at the other cases in our
14 briefing, and what you will see is that the question is whether
15 a reasonably prudent person in the putative plaintiffs'
16 position would have had a reasonable cause to investigate.

17 That is what the standard is.

18 THE COURT: Give me your best case. There are,
19 indeed, cases like that. Give me what you think is your best
20 case. You don't have to be limited to one, but...

21 MS. HURST: *Hirsch v. Rehs Galleries*, your Honor,
22 2020 WL 917213.

23 Another one --

24 THE COURT: All right. So your point is, look, The
25 Times knew what we were doing because they wrote about it, and

P1EsTHE1

1 they wrote about it more than three years before, or I think
2 it's December 2023. That's your point.

3 MS. HURST: Exactly.

4 THE COURT: Done.

5 MS. HURST: And what we were doing was publicly
6 disclosed on the internet. It wasn't a secret.

7 That's all I have on that, your Honor.

8 THE COURT: Fine. Anyone else from the defendants?

9 I think that covers it.

10 Plaintiff. Go ahead.

11 MR. CROSBY: Your Honor. Again, Mr. Crosby.

12 So I wouldn't expect that OpenAI would argue -- and I
13 don't think they argued -- that simply knowing that their
14 models were trained on large volumes of text from the internet
15 without more, just that this is large volumes of text for the
16 internet, would be sufficient information for a specific
17 individual rights holder to bring a lawsuit. Because if that
18 were the case, there would be many more lawsuits than the many
19 that we already have.

20 But, your Honor, the cases in this district, in the
21 Southern District of New York, just our best case was the
22 *Hirsch* case. That was a case where the motion to dismiss was
23 denied. There's only one case that is in the briefing where a
24 motion to dismiss on limitations grounds was granted in a
25 copyright case. And that was the *Minden* case.

P1EsTHE1

1 THE COURT: Because, I take it, the cases are saying
2 it is a fact issue.

3 MR. CROSBY: It's a fact issue. They all say it's a
4 fact issue.

5 THE COURT: Seems to me that's what you want to lead
6 with, it's a fact issue.

7 MR. CROSBY: That's exactly where I'm going with this.

8 THE COURT: Go ahead.

9 MR. CROSBY: These motions are always denied because
10 it's a fact issue.

11 THE COURT: What's the one that you said?

12 MR. CROSBY: Well, it's the *Minden* case.

13 THE COURT: What?

14 MR. CROSBY: The *Minden* case, M-i-n-d-e-n.

15 The interesting thing about the *Minden* case, it was a
16 case involving the seller copyright enforcer who filed many,
17 many lawsuits. And the *Minden* case said you're
18 sophisticated -- this was an argument made in the briefing --
19 that because The New York Times is so sophisticated and does
20 enforce its copyrights vigorously, it should be held to a
21 higher standard than litigants normally are.

22 The case law is clear that there is no general duty
23 for a copyright holder to police the internet for
24 infringements. But the *Minden* case, because you're so
25 sophisticated, we are going to essentially charge you with

P1EsTHE1

1 knowledge of something that was posted to the internet.

2 But then the thing is, that the *Grecco* case in the
3 Second Circuit, which came down this last summer and which was
4 submitted as supplemental authority, expressly regretted this
5 sophisticated rights holder argument and reaffirmed that you do
6 not have a general duty to go out there and police the
7 internet.

8 So the only publicly available information prior to
9 the statute of limitations date that has been pointed out, that
10 I'm aware of, is this paper regarding GPT-2 that was posted in
11 this article. And I think it was the GPT-2 model card that
12 identified specific New York Times works.

13 And the interesting thing about that is that, at the
14 time that that was posted, OpenAI was very much presenting
15 itself as being this nonprofit public benefit entity developing
16 artificial intelligence for the benefit of humanity. And its
17 commercial aspirations at that point were unknown.

18 So to go from the fact that a New York Times reporter
19 wrote an article, that they are generally copying text from the
20 internet, at a time when they are representing themselves to be
21 essentially doing this for this public, for this noncommercial,
22 nonprofit purpose, that that would instigate The New York Times
23 to then go scouring the internet for this particular reference
24 to using their works to train.

25 I think that that is something that is very fact-

P1EsTHE1

1 intensive and something that a fact-finder would need to, sort
2 of, consider and weigh the totality of the circumstances and,
3 you know, hidden discovery is ongoing about how far back
4 OpenAI's ambitions to become a for-profit commercial entity go.
5 It's an issue in the *Musk* suit as well.

6 And so I think we need all of those facts for context
7 to say how likely --

8 THE COURT: I'm sorry. Why would you need to know
9 when they decided to be a profit-making organization --

10 MR. CROSBY: Well, because --

11 THE COURT: -- in terms of knowledge of infringement
12 by The Times?

13 MR. CROSBY: So the key issue in infringement is fair
14 use, whether this is a fair use purpose, copying to train these
15 models. And so, if you're in the position of the --

16 THE COURT: Well, as you said, now ultimately I think
17 would be the issue, not now.

18 MR. CROSBY: Well, that's the ultimate issue in
19 deciding whether to bring a lawsuit, right. The first fair use
20 factor includes whether it was a commercial purpose or whether
21 it was for research, nonprofit, etc., etc.

22 And so, if you're aware that somebody is out there in
23 training models for research purposes or training models in a
24 university or something like that, as opposed to --

25 THE COURT: I understand. I understand.

P1EsTHE1

1 MR. CROSBY: -- to build this answer engine we're
2 talking about, it's very different.

3 THE COURT: I understand.

4 All right. Thank you.

5 MR. CROSBY: Thank you.

6 MS. HURST: Just two quick points in response, your
7 Honor, and I...

8 MR. LIEBERMAN: Excuse me, your Honor.

9 THE COURT: Yes.

10 MR. LIEBERMAN: For the Daily News, I just have one
11 quick point.

12 THE COURT: Go ahead.

13 MR. LIEBERMAN: On the issue of inquiry, the only
14 argument --

15 THE COURT: For the record, tell me your name.

16 MR. LIEBERMAN: My name is Steven Lieberman, and I
17 represent the Daily News plaintiffs and I also represent The
18 New York Times.

19 In the briefing that OpenAI submitted, this is Docket
20 No. 82, with respect to the statute of limitations issue, the
21 only argument that OpenAI made was the statement at page nine,
22 footnote 16, which said, Plaintiffs discovered, or with
23 reasonable diligence should have discovered, those activities
24 prior to April 30, 2021. They provided no specific examples
25 about how the Daily News plaintiffs either were aware of or

P1EsTHE1

1 should have been aware of the fact that its content was being
2 copied.

3 THE COURT: Prior to that date.

4 MR. LIEBERMAN: That's correct.

5 THE COURT: OK. Thank you.

6 MS. HURST: Let me just quickly, before we move on,
7 clarify some facts that Mr. Crosby, I assume, inadvertently
8 mischaracterized for the court.

9 So we have two relevant dates here. One is the date
10 that the so-called model card, which publicly lists a number of
11 the sources that the training data was published. That date
12 was 2019.

13 The second date that is at issue is the date that, as
14 he was describing, OpenAI changed its ambition and began the
15 process of converting to a commercial enterprise from a
16 research one.

17 The facts, which are articulated in the brief, is that
18 by the time The New York Times was writing its articles in 2020
19 about OpenAI and its technology, it was already characterizing
20 OpenAI as gearing up to sell a commercial product.

21 So this suggestion he made that The Times didn't know
22 or couldn't have known at the time that this was all in the
23 service of something other than a research project is simply
24 belied by the facts. Again, I don't want to weight too far
25 into the details here. The fact is, it's not that complicated.

P1EsTHE1

1 There are many disputed facts, but I take your Honor's point if
2 you prefer to reserve that issue for later in the case.

3 THE COURT: All right. Thank you.

4 Again, the fair use issue is for later in the case. I
5 understand what Mr. Crosby was saying is how the first factor
6 is important, therefore, when you switched, as it were, from
7 nonprofit to profit is relevant to the notice.

8 But on this issue, the only thing I need to determine
9 here is whether there are issues of fact. That's it. You're
10 saying no, it's clear. They are saying it's not, let's move
11 on.

12 MS. HURST: Let's move on. You understand the
13 position.

14 THE COURT: I understand the position.

15 MS. HURST: Conveniently, I'll be arguing the
16 contributory infringement point.

17 THE COURT: Well, let's talk about that right now.

18 The idea is that you've contributed, defendants have
19 contributed to infringement by building the LLMs, and enabling
20 the output of the plaintiff's copyrighted material.

21 That's the theory. Speak to me.

22 (Continued on next page)

P1EAThe2

1 MR. GASS: That's right. So we're talking about
2 infringements that are supposedly committed by third parties.
3 So it's a user who takes this --

4 THE COURT: The complaint is that you, OpenAI,
5 ChatGPT, Microsoft, contributed to the copyright infringement
6 of the person who is acquiring these machines and spitting out,
7 according to the plaintiffs, The New York Times articles.

8 MR. GASS: Of course.

9 THE COURT: They say that's copyright infringement and
10 the defendants are enabling the end user to infringe the
11 plaintiffs' copyrights. That's the point.

12 MR. GASS: That's the point.

13 THE COURT: I said that first.

14 MR. GASS: And, your Honor, the legal question whether
15 a technology service provider, like OpenAI, can be secondarily
16 liable for the infringements of those users is resolved by
17 Supreme Court precedent. The *Grokster* case that the Supreme
18 Court decided establishes a crystal clear framework for how you
19 think about when a technology company is going to be
20 responsible for the bad things that its users might do with the
21 tool it's making available.

22 And I'm just going to read you the relevant passage
23 from *Grokster*. This is at 535 U.S. at 942 to 933. The Supreme
24 Court says: In sum, when an article is good for nothing else
25 but infringement, there is no legitimate public interest in its

P1EAThe2

1 unlicensed availability, and there is no injustice in presuming
2 or imputing, [to the technology company] an intent to infringe.

3 Conversely, the Supreme Court says, the doctrine
4 absolves the equivocal conduct of selling an item with
5 substantial lawful, as well as unlawful uses, and limits
6 liability to instances of more acute fault than the mere
7 understanding that some of one's product will be misused.

8 And the Court then goes on at page 937 to expressly
9 hold mere knowledge --

10 THE COURT: I'm sorry. Just a moment. I've lost this
11 feed.

12 (Pause)

13 THE COURT: That quote again, sir. Sorry.

14 MR. GASS: Here was the first quote, and then I will
15 supplement it with a second quote, and I apologize for
16 belaboring the point, but I think the Supreme Court language
17 resolves this issue, so I want to be crystal clear.

18 The first quote which is at 545 U.S. at 932 to 33
19 reads as follows: In sum where an article is "good for nothing
20 else" but infringement, there is no legitimate public interest
21 in its unlicensed availability and there is no injustice in
22 presuming or imputing an intent to infringe. That is, an
23 intent to infringe by the technology company.

24 Back to the quote: Conversely, the doctrine absolves
25 the equivocal conduct of selling an item with substantial

P1EAThe2

1 lawful as well as unlawful uses, and limits liability to
2 instances of more acute fault than the mere understanding that
3 some of one's products will be misused.

4 The Court then goes on several pages later at page 937
5 of the opinion to expressly hold: Mere knowledge of infringing
6 potential or of actual infringing uses, would not be enough
7 here to subject a distributor to liability.

8 Faithfully following the teaching of the Court in
9 *Grokster*, every court to have confronted the question, thought
10 about it and addressed it, has concluded that where you have a
11 technology like the one that's at issue here, it does not
12 suffice to impose secondary liability on the distributor of the
13 technology or service to say they know people are out there
14 misusing it to infringe.

15 That's the holding of the *Luvdarts* case in the Ninth
16 Circuit. It's the holding of *BMG* in the Fourth Circuit. None
17 of the SDNY district court cases that the plaintiffs invoke for
18 a contrary rule, that is, that you could have secondary
19 liability for a technology company on the basis of a general
20 awareness that infringement may happen from misuses, not a
21 single one of those cases deals with a technology like the one
22 at issue here, that everyone agrees is capable of a vast array
23 of non-infringing uses and is not primarily used to infringe.

24 THE COURT: I think I understand the point.

25 MR. GASS: You understand the point.

P1EAThe2

1 THE COURT: Yes. Let's go to the plaintiff.

2 MR. CROSBY: Your Honor, that was -- with apologies,
3 deeply confused.

4 There are two varieties of contributory infringement.
5 One is contributory infringement by inducement, which is
6 encouraging others to infringe. And the other is contributory
7 infringement by material contribution, which is providing
8 assistance to those to infringe knowing that the infringement
9 is occurring, and failing to withdraw that assistance when
10 you're capable of doing so.

11 Sony was about whether you could infer intent to
12 induce infringement solely from the fact of selling a product
13 that had other uses besides infringement.

14 And Sony said that it borrowed the, what's called the
15 Staple Article of Commerce Doctrine from patent law that says,
16 when you sell something that has a substantial non-infringing
17 use, you are not a contributory infringer in patents context.
18 And in that circumstance, they said that civil actually
19 wouldn't make you a contributory infringer in the patent
20 context and it's also not going to make you a contributory
21 infringer in the copyright context.

22 THE COURT: But that's his point.

23 MR. CROSBY: But the point is Sony very specifically
24 said this is limited to the circumstance where the only thing
25 that you do is that you sell this product.

P1EAThe2

1 But in the *Grokster* case, there's some quotes that I
2 would like to continue. So *Grokster* said, and *Sony* also said
3 too, this does not apply in a circumstance, whereas here, you
4 are supplying a product and you're continually operating as a
5 service. You're overseeing it. You have continuous -- you
6 have different interactions. You have more interactions. Or
7 there's other evidence out there from which we could determine
8 your knowledge of infringement or your intent to infringe.

9 All *Sony* says, and it's very clear about this, is just
10 if the only thing you're complaining about is the sale of the
11 product, then substantial non-infringing use negates
12 contributory infringement claim. But if there's other evidence
13 that you knew the product was being used to infringe that you
14 encouraged people to user infringement or that you had
15 continuous interaction with the product while people weren't
16 using it to infringe and didn't, all those things are outside
17 the ambit of *Sony* by *Sony*'s own terms.

18 And so in *Grokster*, the Second Circuit after the
19 portion that was quoted said: We agree with MGM that the Court
20 of appeals misapplied *Sony*, which it read as --

21 THE COURT: When you read, you have to go slower.

22 MR. CROSBY: I'm sorry.

23 We agree with MGM that the Court of appeals misapplied
24 *Sony*, which it read as eliminating secondary liability, quite
25 beyond the circumstances to which the case applied. *Sony*

P1EAThe2

1 barred secondary liability based on presuming or imputing
2 intent to cause infringement solely from the design or
3 distribution of a product capable of substantial lawful use,
4 which the distributor knows is in fact used for infringement.

5 This view of *Sony*, however, was error, converting the
6 case from one about liability resting on imputed intent to one
7 about liability on any theory. Because *Sony* did not displace
8 other theories of secondary liability, i.e. material
9 contribution, and because we find below that it was error to
10 grant summary judgment to the companies on MGM's inducement
11 claim, we do not revisit *Sony* further, as MGM requests, to add
12 a more quantified description.

13 THE COURT: More slowly.

14 MR. CROSBY: On the point of balance, between
15 protection and commerce, when liability rests solely on
16 distribution with knowledge that unlawful use will occur.

17 This issue was also, and the distinction between these
18 cases, these LLM cases, where instead of just a VCR, right, we
19 are essentially loading every movie that was ever made onto the
20 VCR and there is some magical combination of buttons on the VCR
21 that will cause it to reproduce at least some of those movies.
22 Very different from the *Sony* circumstances.

23 Judge Tigar in the *Andersen* case in denying a motion
24 to dismiss on this very ground said: The theory of this case
25 is not similar to, for example, a case asserting contributory

P1EAThe2

1 infringement based on the sale of VCRs where, after discovery,
2 plaintiff had no evidence of defendant's intent to induce
3 infringement.

4 The Supreme Court explained that in those
5 circumstances, intent could not be based on presuming or
6 imputing intent to cause infringement solely from the design or
7 distribution of a product capable of substantial lawful use,
8 which the distributor knows is in fact used for infringement.

9 Instead, this is a case where plaintiffs allege that
10 Stable Diffusion -- the AI model producer in that case -- is
11 built to a significant extent on copyrighted works and that the
12 way the product operates necessarily invokes copies or
13 protected elements of those works. The plausible inferences at
14 this juncture are that Stable Diffusion by operation by end
15 users creates copyright infringement and was created to
16 facilitate that infringement by design. In addition to the
17 comment of Stability's CEO, plaintiffs reference articles by
18 academics and others that training images can sometimes be
19 reproduced as outputs from the AI products.

20 And we have also cited articles about users, in while
21 using these products, to evade paywalls. These are not just
22 examples that we curated and created ourselves. For example,
23 there is a -- this is cited in the brief and it's also cited in
24 the Daily News complaint and could be specifically cited in The
25 New York Times complaint if amendment were required, that I

P1EAThe2

1 think demonstrates the plausibility of the allegations,
2 reporting that users were posting to Reddit forms, to social
3 media on the internet, examples of how they had gotten around
4 the paywall using ChatGPT, and using a product called
5 SearchGPT. And, in fact, OpenAI pulled the product after those
6 reports came out. So they were aware that these products were
7 being used to infringe.

8 So all of this information goes far beyond the bounds
9 of *Sony* and is exactly on point with *Andersen* case.

10 THE COURT: Thank you. I think I understand.

11 MR. CROSBY: There's only one more point I really want
12 to make here, your Honor. And I think you pointed at Count
13 Four. And all we've been talking about is secondary liability
14 for outputs that are in response to user prompts. And I just
15 want to make clear, we have a second contributory infringement
16 count that is against Microsoft alone. I think it's Count
17 Three. And that is asserting that Microsoft, to the extent
18 that they are not directly liable for the copying that was done
19 to train the GPT models, is liable as a contributory infringer
20 because they knew that OpenAI was using copyrighted works
21 without permission to train those models, and provided -- among
22 assistance, including computing infrastructure, technical
23 support, etc., etc., knowing that that infringement was going
24 on for the purpose of fostering that infringement. And that is
25 not something that's been challenged or addressed in any of the

P1EAThe2

1 briefing or the argument that I've heard so far.

2 THE COURT: How does that differ from the prior
3 arguments?

4 MR. CROSBY: So the prior argument all concerns
5 secondary liability for infringing outputs that are generated
6 in response to user prompts when the models are deployed.

7 THE COURT: Right.

8 MR. CROSBY: And this is secondary liability for the
9 copying that incurs at ingestion when the models are trained.

10 THE COURT: Okay.

11 MR. CROSBY: And that I don't think that has been
12 joined at all in the briefing. That's Count Three, your Honor.

13 THE COURT: Okay. Sir?

14 MR. LIEBERMAN: Your Honor, Steve Lieberman.
15 Directing your Honor's attention to paragraph 147 of the Daily
16 News complaint. In support of Mr. Crosby's argument, OpenAI --

17 THE COURT: Let me just look at it.

18 Yes, sir, go ahead.

19 MR. LIEBERMAN: OpenAI has something called a custom
20 GPT Store.

21 THE COURT: I saw where you can -- it's designed to
22 have you avoid the payroll.

23 MR. LIEBERMAN: Exactly right, your Honor, the removed
24 payroll product. And the news summarized the product, which is
25 described as encouraging users to save on subscription costs

P1EAThe2

1 and to skip payrolls using the link, text, or URL.

2 That point supports Mr. Crosby's argument. And I
3 would suggest the proper analysis here is found in your Honor's
4 own decision in the *Arista v. MP3* case, which was also a
5 material contribution issue, where your Honor denied summary
6 judgment by MP3 on very similar grounds.

7 THE COURT: All right. I always find the ones that
8 are summary judgment cases not that helpful when we're really
9 talking motion to dismiss. But everybody is trying to use
10 them.

11 What do you do with 147?

12 MR. GASS: Your Honor, that falls squarely in the
13 category of uses that the Supreme Court said in *Grokster* would
14 not be enough to subject the technology distributor to
15 secondary liability. Mere knowledge of infringing potential or
16 of actual infringing uses. There is no allegation in paragraph
17 147 that OpenAI had any idea that someone, some third party,
18 had created this custom GPT, much less that OpenAI knew about
19 it and then did nothing to prevent it.

20 This is a user apparently engaging in an actual or
21 potential infringing use. And there is nothing in the record
22 to suggest any particularized knowledge of OpenAI of those uses
23 at a time that OpenAI could do anything about it. That is the
24 quintessential circumstance where the provider of the
25 technology service is not liable for the infringements of its

P1EAThe2

1 users.

2 THE COURT: Okay.

3 MR. GASS: On that point, and four very quick points.

4 One, you heard Mr. Crosby say that the core
5 distinction here is between selling a product and offering a
6 service. The Fourth Circuit squarely and expressly rejected
7 that distinction in the *BMG* case I'll give you the cite for
8 that. 881 F.3d 293 at 308.

9 Second, on the *Andersen* opinion, once again,
10 Mr. Crosby mischaracterized it a little bit. Here, I'm even
11 more confident he did so inadvertently. It was not Judge
12 Tigar. It was Judge Orrick. I know this because that is one
13 of my cases. I am representing one of the defendants in that
14 case. And the key passage from the motion to dismiss
15 opinion --

16 THE COURT: Are you talking about *Andersen* against
17 Stability?

18 MR. GASS: Yes. Exactly. The key passage from that
19 opinion is the part where Judge Orrick says that: As he
20 interprets the allegations in the complaint, all of the outputs
21 "necessarily invoke copies." So that, in the paradigm of the
22 Supreme Court, puts this outside of the cases that we've been
23 talking about here today that are all about technologies that
24 do not invariably, or to use Judge Orrick's term, necessarily
25 contain infringements in their outputs.

P1EAThe2

1 THE COURT: Why don't we conclude this point.

2 MR. GASS: One more critical point, your Honor, and
3 then I will yield the floor to Ms. Hurst, who I think wants to
4 talk about the Microsoft point.

5 This is the most important point. I agree with
6 Mr. Crosby that the cases draw a distinction between technology
7 company secondary liability situations where the theory is that
8 they've induced infringement, and technology company secondary
9 liability theories where there's not an inducement claim.

10 The problem for Mr. Crosby is that none of the
11 plaintiffs here have alleged or could, consistent with their
12 Rule 11 obligations in good faith, allege that OpenAI has been
13 out there actively promoting people to use their tool to engage
14 in infringement. That is the standard from *Columbia v. Fung*,
15 710 F.3d 1020. Inducement means that it has to be -- the
16 technology company has to have an object of promoting the use
17 of the tool to infringe. No one could say that that is
18 happening here. If you search these complaints for the word
19 induce or inducement, you won't see it anywhere close to the
20 discussions of secondary liability.

21 I will leave it at that. Ms. Hurst.

22 MS. HURST: Your Honor, briefly. I think Mr. Crosby's
23 reference to Count Three against Microsoft is an excellent
24 comparison tool here for why Count Four does not meet the
25 standard.

P1EAThe2

1 In Count Three, The New York Times in paragraph 176
2 alleged that Microsoft directly assisted OpenAI in copying the
3 training datasets containing the copies of The Times works.
4 Now, your Honor, that's not true, but at the pleading stage, we
5 know we have to accept it's true. That allegation shows actual
6 knowledge of specific acts of alleged infringement.

7 There's nothing like that in Count Four, your Honor.
8 That's a specific allegation of copying their works and actual
9 knowledge on the part of Microsoft. But there's nothing like
10 that in Count Four with respect to the operation of these
11 models. There's no example of end user infringement that
12 either defendant is alleged to have knowledge of. There's no
13 allegation of end user infringement that Microsoft is alleged
14 to have had knowledge of.

15 And, your Honor, that leaves only this purported
16 device versus service distinction, but we know that's not the
17 law, your Honor. We know it's not true because *BMG* and
18 *Luvdarts*. We also know it's not true because of *Cartoon*
19 *Network* in this circuit. And to use the exact example posited
20 by Mr. Crosby, the next generation of VCRs, which is a DVR
21 service, in both *Cartoon Network* and *Fox v. Dish*, those were
22 fair uses. There was no end -- without specific acts of end
23 user infringement, there could be no secondary liability, your
24 Honor.

25 THE COURT: All right. Let me turn Mr. Crosby then on

P1EAThe2

1 the question of specific -- the need, alleged need, to have
2 specific examples set forth in order to sustain the complaint.

3 MR. CROSBY: Well, your Honor, that's just not
4 correct. I mean, so all of the cases that have been cited on
5 this issue are all summary judgment cases. So they're all
6 cases that got past the pleadings where we looked at was there
7 evidence of knowledge at the time that we finally got
8 discovery. None of these were dismissed on the pleadings
9 because you didn't sufficiently allege knowledge.

10 So the *Matthew Bender* case, which is one they've
11 leaned on heavily, that was a case where at the pleading stage,
12 the only --

13 THE COURT: Is that the one where the only examples
14 were done by the lawyer?

15 MR. CROSBY: Right. And it was an interesting case.
16 So it was Westlaw. So they have a thin copyright or no
17 copyright on most of the materials in their casebook. But they
18 said, you know, somebody could essentially reverse engineer and
19 create one of our casebooks, you know, and that therefore,
20 infringe the compilation authorship or thin authorship in the
21 organization of that from using this CD ROM product that uses
22 all these casebooks.

23 And so at the pleading stage, all there was, was this
24 example of the lawyer having successfully done that. And they
25 got through discovery and it was determined that in fact nobody

P1EAThe2

1 had ever done that. It was probably implausible that anybody
2 would ever do that.

3 This is a very different case. It is widely known
4 that these models are capable with appropriate prompting of
5 regurgitating. And we've shown they can be prompted to
6 regurgitate materials that have essentially been imprinted in
7 the parameters of their model. And that's, by the way, the
8 definition of a copy under 17, U.S.C., Section 11. If
9 there's -- by any means you can perceive a work from something
10 that it's been -- from a fixation. So if there's a magical
11 prompt that would allow you to recall a work, that means that
12 there's a copy in the model.

13 And so we know, and it's been widely reported and we
14 have alleged, that this is a common reproducible phenomenon.
15 And the New York Times and Daily News and journalism in general
16 are highly represented in the training sets of these models and
17 are highly likely to be memorized.

18 So it is certainly plausible in view of those facts
19 that a user who is trying to get around a paywall or trying to
20 just find out something that the New York Times reported and
21 doesn't have a subscription or doesn't want to go to the New
22 York Times could in fact do this. And in discovery we will
23 find out whether this happened or not, as in *Matthew Bender*,
24 and how often this happened or not.

25 And the argument that, well, we're going to open the

P1EAThe2

1 flood gates of discovery into outputs if we allow this claim to
2 go forward, there's no flood gates. The discovery has already
3 happened. And the reason is this is a claim in the
4 alternative. So, to be clear, we claim that Microsoft and
5 OpenAI are direct infringers when a user prompts the model to
6 output copyrighted content.

7 And the reason is, as in *Cartoon Network* and as
8 Justice Scalia's concurrence in the *Aereo* area case, and all
9 the circuits that have since considered it, the key test for
10 who is a direct infringer, is who selects the works that is
11 ultimately copied. So the cases like *BMG*, these service
12 provider cases, these are cases where you have essentially a
13 service that takes content that is posted by users and then
14 gives content to users. You're just sort of an intermediary.
15 You're not selecting the content, and those circumstances to be
16 secondarily liable.

17 The Ninth Circuit has a test where if you have
18 knowledge of specific infringements, and you have simple means
19 available to protect those infringements if you are essentially
20 a passive service provider, that's the requirement to be a
21 contributory infringer. Right? But if you're actually
22 selecting the works and essentially making them available for
23 users to elicit, that makes you a direct infringer. They're
24 probably going to disagree with that.

25 So this is a claim in the alternative that even if

P1EAThe2

1 they're not deemed to be a direct infringer, when a user
2 prompts an infringing work, that they're still contributorily
3 liable because they selected -- because they provisioned the
4 service with the works, either at training or by giving it the
5 Bing search index. So they know which works are used, which
6 works are present in the search index, which works are present
7 in the training sets, and then they monitor and log both the
8 prompts and the outputs whenever somebody interacts with it.
9 So they know which works are output.

10 So we are going to discover, and we are already in the
11 process of discovering, all the outputs of these models that
12 they have reported in their logs for purposes of our direct
13 infringement claim. And if the direct infringement claim
14 fails, then we will use the same evidence to support our
15 contributory infringement claim. But there's to discovery of
16 output that is going to occur as a result of this claim being
17 allowed versus being denied.

18 THE COURT: I understand.

19 MR. CROSBY: And we should view the summary judgment
20 like all the cases.

21 THE COURT: I understand.

22 MR. GASS: Last rejoinder, not like all of the cases
23 because *Luvdarts v. AT&T Mobility* addresses this properly on a
24 motion to dismiss.

25 THE COURT: Thank you. Let's go to the DMCA claim.

P1EAThe2

1 And we need to pick up the pace a bit. Let's start with the
2 issue of removal of the CMI in the training process. Go ahead,
3 sir.

4 MR. GRATZ: Joe Gratz for OpenAI. With respect to
5 the --

6 THE COURT: We'll talk about the removal of CMI in the
7 training process.

8 Go ahead. Give me yours first.

9 MR. GRATZ: So in the interest of picking up the pace,
10 your Honor, I want to focus with respect to the 1202 claim on
11 what we regard as sort of their biggest problem. And their
12 biggest problem is they don't have a plausible story for how
13 they would be better off if the CMI they say was removed was in
14 fact removed.

15 THE COURT: Is that your injury point?

16 MR. GRATZ: Correct.

17 THE COURT: Okay.

18 MR. GRATZ: And that problem plays out in a number of
19 ways, and has a number of consequences. It means they don't
20 plead facts showing they were injured as the statute requires,
21 as 1203 requires. It means they don't plead facts showing they
22 have an injury in fact, as Article III requires, and that same
23 kernel of problem plays out in the fact that they don't plead
24 facts showing how anybody, particularly us, would have any
25 reason to think in these circumstances that the absence of CMI

P1EAThe2

1 induced, enabled, facilitated, or concealed any infringement.

2 The problem they have is that there is no -- they do
3 not have a coherent story of causation.

4 THE COURT: I'm sorry. Isn't your basic point here of
5 they don't show any injury by the removal of the CMI?

6 MR. GRATZ: That's right. There is not a way in which
7 the world would be better for them in the ways that they say
8 the world is not good for them if the CMI that they say was
9 removed was never removed.

10 THE COURT: Okay. I understand that point. What else
11 did you want to tell me before I turn to them?

12 MR. GRATZ: So that is, that is the key point.

13 THE COURT: And that's what I thought.

14 MR. GRATZ: I will address whatever they come up with
15 as to how the world would be different for them if the CMI they
16 say was removed --

17 THE COURT: Well, the way I phrase it is: What's the
18 injury? Sir? Or did you want to say something? Go ahead.

19 MS. HURST: Your Honor, just briefly on behalf of
20 Microsoft. There's no allegation that Microsoft did anything
21 that's tantamount to a removal. So in addition to the injury
22 argument, there's simply a paucity of any allegation that
23 Microsoft committed the alleged act that violates the statute.

24 THE COURT: All right. Let's turn to the plaintiff.
25 Sir?

P1EAThe2

1 MR. LIEBERMAN: Your Honor, Steve Lieberman.

2 Let me start first with the question that you had
3 asked defendants that they didn't answer, and that's the
4 removal. If we look at the Daily News complaint at paragraphs
5 161 and 162, there were very specific and detailed allegations
6 about two different programs that OpenAI used to remove CMI.
7 There are what are called content extractors so they do exactly
8 what the name suggests.

9 THE COURT: Let me read those. Just a moment.

10 Purposeful removal of CMI during the training; isn't
11 that what that is?

12 MR. LIEBERMAN: It is, your Honor.

13 THE COURT: Got it.

14 MR. LIEBERMAN: So there are two programs, one called
15 Dragonite and one called Newspaper. Newspaper actually has a
16 function called trim. What these programs do is they remove
17 the content and they separate it from the CMI: The name of the
18 author, the title of the article, the date, the links, the
19 copyright notice, etc.

20 This is adequately pled. I want to come back to the
21 Microsoft issue, but let me turn next to your Honor's very
22 specific question that the harm and whether the harm is
23 cognizable.

24 THE COURT: Don't I need to find that?

25 MR. LIEBERMAN: Don't you need to find harm, you do

P1EAThe2

1 need to find harm sufficient to establish Constitutional
2 standing. The statute, if you have Constitutional standing,
3 the statute provides statutory standing. And if we can look at
4 the statute, your Honor, first maybe we start with that.

5 There are two 1202(b) claims that the New York Times
6 and Daily News plaintiffs are asserting. One is 1202(b)(1),
7 which is intentionally removing or altering any copyright
8 management information. It doesn't require distribution. It's
9 intentionally removing or altering any copyright management
10 information, knowing or with respect to civil remedies under
11 1203, which is what we're dealing with here, having reasonable
12 grounds to know it will induce, enable, facilitate, or conceal
13 an infringement of any right under this title.

14 So (b)(1) doesn't have to do with distribution. It
15 has to do with the removal itself. And the question of (b)(1)
16 removal of CMI was dealt with by the Second Circuit in the
17 *Mango* case. And if I can just read briefly from what the
18 Second Circuit said about the statute.

19 The Second Circuit said: Fearful that the ease with
20 which pirates could copy and distribute a copyrightable work in
21 digital form was overwhelming the capacity of conventional
22 copyright enforcement to find and enjoin unlawfully copied
23 material, Congress sought to combat copyright piracy in its
24 earlier stages before the work was even copied.

25 Before the work was even copied. That is, the Second

P1EAThe2

1 Circuit says what Congress intended to do with this statute was
2 to prevent people from removing the CMI before anything else
3 happened. So the non-digital annuli would be you go into a
4 research liability and find somebody's senior thesis, and you
5 take an X-Acto Knife. Your Honor probably remembers X-Acto
6 Knives, and you --

7 THE COURT: And I remember senior theses, too.

8 MR. LIEBERMAN: As I do as well. You go into
9 Firestone library, you find the senior thesis, you cut it out
10 with an X-Acto Knife, doesn't have the author's name in it. Or
11 maybe you alter it by putting in somebody else's name. It's
12 sitting there in the library.

13 So the question: What's the harm? What the Second
14 Circuit said in *Mango* was let's look to see what Congress'
15 purpose here was. And Congress' purpose was to try to stop
16 what it was afraid it was going to be massive digital piracy.
17 And to look -- I would like to direct your Honor's attention to
18 the legislative history on the statute, which is cited in the
19 Daily News brief at page 16.

20 The Senate report at page 8 said: Due to the ease
21 with which digital works can be copied and distributed on the
22 internet, "virtually instantaneously copyright owners will
23 hesitate to make their work available on the internet without
24 reasonable assurance that they will be protected against
25 massive piracy."

P1EAThe2

1 So here's the harm -- and also before I jump into the
2 harm.

3 THE COURT: Where is the harm? That's what I want to
4 know.

5 MR. LIEBERMAN: Okay.

6 THE COURT: Starting with where's the harm.

7 MR. LIEBERMAN: The harm is the training databases
8 that OpenAI reposed the content in after extracting the CMI are
9 not hermetically sealed containers. They're not closed
10 databases. These are the databases -- this is the database,
11 this is the model that people are going to query. They're
12 going to ask questions of it. They may ask questions about the
13 Super Bowl in 2003. They may ask questions about a
14 presidential debate from 2010. I'm going to get separately to
15 Microsoft and the more recent stuff. And that when they do
16 that, what will happen is because the CMI has been extracted,
17 because the CMI has been extracted, what you're going to get
18 out as an output is you're going to get either the verbatim
19 language or a summary of a New York Times or a Chicago Tribune
20 or Daily News article without the attribution to the New York
21 Times or the Daily News.

22 THE COURT: What's the injury? I understand that.
23 That's stripping the CMI. What's the injury?

24 MR. LIEBERMAN: This is exactly the injury, your
25 Honor, that the Second Circuit was talking about in describing

P1EAThe2

1 the statute. You're leaving people open for massive copyright
2 infringement without the ability to trace it. And here, let me
3 give you two examples of specific injury.

4 First, if you look at the Daily News complaint in
5 paragraph 167. What you see -- I'm sorry. 168, 169, your
6 Honor. If you look at 168, the query to the Daily News -- the
7 query to the accused models to Copilot was: I need content for
8 my blog, please provide a news article about what the Mets see
9 in Julio Teherán.

10 And then you get an answer from Copilot. And at the
11 end of the answer it says: Feel free to incorporate this
12 information into your blog. No cite to the Daily News. It
13 doesn't say this comes from the Daily News. It doesn't say
14 this was a Daily News article. In fact, at the bottom there's
15 a little line that says learn more, and it gives four different
16 sources. Three of them are not the Daily News.

17 This is what the removal of the CMI almost inevitably
18 leads to. It's like it causes the alarm system in your house
19 to go down.

20 The Second Circuit in *Mango* said, when you're looking
21 at the next step, your Honor, when you're looking at the step
22 of scienter, the Second Circuit said that that scienter can be
23 satisfied if you have a reasonable basis to believe that you're
24 hiding your own infringement. So what happens here? OpenAI
25 takes Daily News or New York Times content. They strip out the

P1EAThe2

1 CMI. How does one find that this content that they've copied,
2 that we contend is a copyright infringement, came from The
3 Times or came from the Daily News.

4 In this case, we have had to have a team of experts
5 spend literally weeks, hundreds and hundreds of hours, doing
6 searches in their models using what are called URLs, and Ngram
7 searches to try to locate the publishers' content.
8 Fortunately, we found millions of examples of verbatim content
9 in their models. But we would not have had to have done this
10 if they didn't take steps to conceal their own infringement,
11 which is a violation, concealing your own infringement is a
12 violation under *Mango*. And we have a case directly on point on
13 this point, your Honor, on the standing issue, on the harm.
14 It's the furniture net -- *FurnitureDealer.net* case, which is
15 cited in our briefs.

16 That was a case -- well, here's the language from the
17 Court ruling: Plaintiff has met the requirements of an injured
18 person based solely on their allegation that defendant's
19 removal of CMI required plaintiff to undertake an investigation
20 to discover where its content was appearing online. Had
21 defendants not removed CMI, plaintiff could have avoided
22 undertaking these efforts to uncover the true source and the
23 full extent of the alleged infringement.

24 So you have got a variety of different harms. One
25 is --

P1EAThe2

1 THE COURT: Well, your argument is it leaves open the
2 possibility of massive wrongdoing?

3 MR. LIEBERMAN: Which is exactly what the purpose of
4 the statute was, your Honor.

5 THE COURT: All right. I understand that.

6 MR. LIEBERMAN: And the Supreme Court in the
7 *TransUnion* case said that when you're looking at the issue of
8 harm and standing, it's instructive to see what Congress, in
9 passing the legislation, intended to be the harm that it was
10 addressing. And here, it's very clear that the statute, as the
11 Second Circuit has interpreted, has made clear the harm is
12 something which occurs right at the very beginning when the CMI
13 is removed. It doesn't require --

14 THE COURT: I understand the point.

15 MR. LIEBERMAN: Thank you.

16 THE COURT: Is there anything else you wanted?

17 MR. LIEBERMAN: Just one specific example, your Honor.

18 No. I can wait for further questions on this issue,
19 your Honor.

20 THE COURT: All right. Fine yes, sir.

21 MR. TOPIC: Hi, your Honor. I have a few additional
22 comments on behalf of CIR. We agree with what was just said,
23 but I think we have maybe sort of a complementary or additional
24 take on this.

25 It's the take that we argued in the Intercept case,

P1EAThe2

which recently survived dismissal on similar allegations. To answer your question of what is the injury, the injury is the copying. The requirement under *TransUnion* is simply to analogize to historically recognize injury. A copyright law has always, always required no more than copying. It doesn't require any harm beyond that. It doesn't require any dissemination. It doesn't require any economic loss. The copying is the harm.

TransUnion recognizes the difference between an injury in law and an injury in fact. And so the fact that copying may not be an element of a DMCA claim, is the same as the dissemination wasn't an element of the claim in *TransUnion*. It's part and parcel of the conduct that can be remedied by the case or by the claim. And therefore, the analogy to the mere act of copying is sufficient. And we have alleged that in the course of creating their training sets, in the course of removing the CMI, they made copies of our works. They didn't have properly authorized copies that they just took the proverbial X-Acto Knife to. They made copies and omitted the CMI.

So the particular way in which the violations occurred here are not only analogous. They're identical to a copyright infringement claim. That is a far closer analogy than what the Second Circuit allowed in the *Saba* case, which analogized harm to voting rights for certain types of shares to trespass to

P1EAThe2

1 chattel. Those are different things. There are all types of
2 elements to overlap. We have a much tighter analogy between
3 DMCA -- between the injury here and the injury under copyright
4 because they're exactly the same thing.

5 We also have another theory of harm or of injury
6 sufficient for standing, and that's unlawful profits and unjust
7 enrichment.

8 THE COURT: Unlawful profits?

9 MR. TOPIC: Or unjust enrichment. You can think of it
10 the same way. Under DMCA, one of your remedies is to disgorge
11 the profits that were derived by the plaintiff -- or by the
12 defendant as a result of the violations. And under Donahue and
13 Packer, which we cite in our brief: A defendant's profits
14 resulting from -- I'm paraphrasing or synthesizing.
15 Defendant's profits resulting from a statutory violation that
16 provides for profit disgorgement are considered a plaintiff's
17 injury where the statute protects that plaintiff from the
18 conduct resulting in the profits.

19 So, in other words, their unjust enrichment is our
20 injury sufficient for standing because there's no question that
21 disgorgement of profits has always been a historically
22 recognized injury that can be recognized in unjust enrichment
23 law or under copyright law for that matter.

24 The last point I want to make is while they say we
25 have to disseminate works in order to have a claim, which is

P1EAThe2

1 coming from *TransUnion* only because the plaintiff in *TransUnion*
 2 analogized to defamation, which we don't do here. But let's
 3 set that aside. But even if we were required to show
 4 dissemination, we have alleged, we have (b) (3) claims arguing
 5 that they have shared the copies back and forth in the course
 6 of their partnership in the course of Microsoft providing --

7 THE COURT: That's distribution point under (b) (3)?

8 MR. TOPIC: Correct.

9 THE COURT: Okay.

10 MR. TOPIC: Correct. So to the point there is a
 11 requirement of dissemination, as the plaintiffs claim, even
 12 though there is not, we have alleged that there is
 13 dissemination between the two defendants.

14 Last thing I'll say, because we touched a little bit
 15 on the scienter requirement, is the very liberal pleading
 16 standards for scienter in a DMCA case. We cited *Hersch*, where
 17 the allegations were described as sparse. It was the removal
 18 of gutter credit in a photo. So these are a fact -- scienter
 19 is not suitable for a motion to dismiss. What they knew, all
 20 the elements for scienter, those are the kind of things that
 21 are more appropriate for trial or for summary judgment.

22 MR. LIEBERMAN: And, your Honor, I realize I neglected
 23 to negligent Ms. Hertz point. Just one sentence on that.

24 THE COURT: Go ahead.

25 MR. LIEBERMAN: I direct the Court's attention to

P1EAThe2

1 paragraph 168 of the Daily News complaint. That's a Microsoft
2 Copilot output. The content had CMI in it. Microsoft, we
3 believe and we allege, stripped out the CMI, and that's why the
4 output doesn't have any reference to the Daily News.

5 Number two, your Honor made the point a little earlier
6 about harm resulting from lawyers putting in inquiries
7 themselves. That issue is addressed in the standing context
8 just recently by Judge Tigar in the *Doe 1 v. GitHub* decision.
9 This is the 2024 WL235217 at pages 4 and 5. There, Judge Tigar
10 held that plaintiffs have standing to assert the damages claim
11 under the DCMA Sections 1202(b)(1) and (b)(3) based on the
12 output of code in response to their own prompts. That is, they
13 did the prompts, they put them in the amended complaint, and he
14 said that was satisfactory to show standing, both
15 Constitutional and statutory standing.

16 THE COURT: Okay.

17 MR. LIEBERMAN: Thank you, your Honor.

18 THE COURT: Let's move on.

19 Common law, unfair competition by misappropriation.
20 Who wants to speak to that?

21 MR. GRATZ: I would be happy to speak to that, your
22 Honor. I feel compelled to say before I begin that we still
23 haven't heard any harm from the absence of CMI.

24 THE COURT: We've been through that. We've minded.
25 You have your point of view. They have their point of view.

P1EAThe2

1 At some point, you will learn my point of view. Go ahead.

2 MR. GRATZ: Thank you, your Honor.

3 So with respect to the misappropriation claims, there
4 are two big problems. The first big problem is that one of the
5 elements that they need to plead is lack of attribution and all
6 of their examples contain attribution of one kind or another.

7 So the first thing I want to talk about, because this
8 is something I think they may dispute, is that lack of
9 attribution is something that is a necessary element to bring a
10 non-preempted hot news misappropriation claim. And this comes
11 from the *Barclays v. Flyonthewall* case, 650 F.3d at 903. And
12 it rejects the claim on a number of bases. But one of the
13 bases on which it rejects the claim is that the defendant was
14 "selling the information with specific attributions to the
15 issuing firm," that is, the plaintiff there. So it could not
16 have been selling the material "as its own" which the Second
17 Circuit says is something that is necessary for a claim like
18 this to survive preemption.

19 The facts there were that they were ascribing the
20 material to its creator. There was proper attribution and they
21 said that where that was true, there couldn't be a claim. And
22 that is true here. That is, in the New York Times' examples,
23 they either all say in words that this is from the New York
24 Times, or there is a footnoted link to the New York Times'
25 website.

P1EAThe2

1 In the Wirecutter examples, it says it's from
2 Wirecutter. In the New York Daily News examples, these are
3 from Copilot, but there is attribution either in words or as a
4 footnote or a link. So in each of these cases --

5 THE COURT: Yeah, but it seems to me, I think the wire
6 cutter point is that there's no link. And they're losing the
7 income if people on the actual wire cutter click on the link.

8 MR. GRATZ: That was -- I'm sorry, your Honor.

9 THE COURT: Just a moment.

10 (Pause)

11 MR. GRATZ: That was precisely the plaintiffs'
12 argument in the *Barclays* case that the Second Circuit said was
13 insufficient. That is, that while there was attribution, you
14 knew where it came from, there was some other benefit that they
15 weren't getting even though there was attribution.

16 So the same is true of the allegations here. It is
17 not being characterized as OpenAI's or Microsoft's
18 recommendation on what toaster to buy. It is being
19 characterized on the Wirecutter's recommendation on what
20 toaster to buy. And that attribution to source eliminates the
21 possibility of un-preempted claim under the *Barclays* case,
22 which follows the Supreme Court's *INS* case.

23 THE COURT: Okay.

24 MR. GRATZ: And that is true of each and every one of
25 their examples. So that is the first of their -- of the

P1EAThe2

1 problems.

2 I think that -- and that is the cross cutting problem
3 that eliminates all of the hot news examples for all of the
4 plaintiffs who have brought those claims.

5 The second big problem is that another necessary
6 element to avoid preemption is that the material is highly time
7 sensitive. And none of their examples are sufficiently time
8 sensitive. The highly time sensitive --

9 THE COURT: Yeah, I do have a question on that. So I
10 do want the plaintiffs to respond to that. Go ahead.

11 MR. GRATZ: This is an element -- this is *FII v.*
12 *Moody's* case in the Second Circuit is the key case here where
13 information that was 10 days old was too stale to support a hot
14 news claim. And here, with respect to OpenAI, all of the
15 information is at least months old. You know, we have things,
16 hot news like what chair to get, which has been the same
17 recommendation since 2017. Some of the things that they are
18 saying --

19 THE COURT: That's not the point. The point is how
20 close did the recommendation being made, is it then copied by
21 you, as they say.

22 MR. GRATZ: So, your Honor, that's actually I think
23 not importantly the question. The question is how time
24 sensitive is this information. How much of the value of this
25 information comes from -- it dissipates very quickly. And that

P1EAThe2

1 was true in the *INS* case. And that was not true in the *FII v.*
2 *Moody's* case because they said, look, this was information that
3 wasn't disseminated for 10 days after it was first disseminated
4 by the plaintiff. And by that -- that's far too long to
5 support a claim like this. As opposed to in cases where this
6 was true, like I think they found it to be true in the *NBA*
7 case, this element was present, that what's the current score
8 of an ongoing baseball -- basketball game, is highly time
9 sensitive because, you know, this information would be no good
10 in the future.

11 THE COURT: Yes. I have the point.

12 MR. GRATZ: So as to -- particularly as to the
13 examples with respect to OpenAI, all of it is months old and
14 there's no plausible argument that it is so time sensitive that
15 it was reproduced before --

16 THE COURT: You made the point. I understand it.

17 MR. GRATZ: Those are my points with respect to this
18 claim, your Honor.

19 THE COURT: All right. Thank you.

20 Sir?

21 MR. LIEBERMAN: Your Honor, Steve Lieberman again.

22 Let me start with point number one and then I'll do
23 point number two. The lack of attribution, we've got two
24 decisions from the Second Circuit setting out what the
25 requirements are in order to fall outside of preemption for a

P1EAThe2

hot news unfair competition claim. There's a three-part test. There's a five-part test. Not clear whether there's any difference between the two of them. But one of the elements of either the three-part test or the five-part test is not that there has to be no attribution. That's not one of the elements.

And, in fact, the *Barclays* case was not decided on that basis. The *Barclays* case was decided on the very clear ground that there was no free riding because what Flyonthewall was doing is it was reporting on the recommendations from the brokerage houses with respect to stocks, and that itself was the news. That is, it wasn't free riding. It was reporting on the news that the brokerage houses had just put a buy order or a buy recommendation on this particular stock. And the Court is very explicit. Judge Sack is very explicit that that is the basis for his decision.

Now, it may well be that every once in a while, there's, you know, the fact that The New York Times or the Chicago Tribune or the Daily News reports on something, that that's the news. But that's really a very small exception. Most of the time what's going on here, as we've alleged, is that they simply reproduce either verbatim or the facts from or a summary of an article. And that is not doing what the Second Circuit said in the *Barclays* case.

In terms of the time sensitive nature, if you look at

P1EAThe2

1 the Daily News complaint at paragraphs 127 to 128, we have the
2 defendants copying a Daily News article the very same day the
3 article appeared. In 118 and 119, you've got a report on a
4 lunar eclipse, that's days. 121 and 122, a triple homicide,
5 that's one day. 124 and 125, it's the Yankees schedule, that's
6 one day. 136 and 137, the Warriors are serging, one day.

7 Imagine this, at least for the Daily News comings, a
8 lot of these papers are papers where the focus of the readers
9 is on what happens in that particular area. Some of them, the
10 New York Daily News, the Chicago Tribune, etc., they have broad
11 interest nationwide. But if you're going to subscribe to a
12 local paper in order to get high school basketball scores and a
13 description of what happened in the game the night before, and
14 you can get that for free just by typing it into ChatGPT and
15 Bing Copilot, that's the free riding that the Second Circuit in
16 both *Barclays* and *NBA* was talking about.

17 Now, with respect to wire cutter, your Honor was
18 exactly right about the problem. You are disconnecting wire
19 cutter right at the point where it is about to reap revenue.

20 THE COURT: Right at the point where?

21 MR. LIEBERMAN: It's about to reap revenue.

22 THE COURT: Oh.

23 MR. LIEBERMAN: It's been a month testing different
24 ovens, different refrigerators. It's done all types of work.
25 It produces its recommendations and its reasons. It's got a

P1EAThe2

1 link. And if you go from that link and you buy the product,
2 wire cutter gets a fee.

3 If OpenAI can copy those recommendations, even if they
4 attribute it to wire cutter, that doesn't diminish the harm to
5 wire cutter. It doesn't diminish the harm to wire cutter
6 because if somebody goes and buys that product and they don't
7 do --

8 THE COURT: They're not going to be doing the link,
9 which is the key to your revenue.

10 MR. LIEBERMAN: Right. Indeed, the attribution may
11 even cause additional harm to The Times.

12 THE COURT: How?

13 MR. LIEBERMAN: Well, we have examples in the
14 complaint where they cite wire cutter as endorsing certain
15 products, which they didn't endorse. Wirecutter's reputation
16 can be very fragile. If they endorse a particular -- if
17 ChatGPT says Wirecutter endorsed a particular product, and that
18 product was, for example, the subject of a product recall or
19 kills somebody's child who buys the product, the harm runs down
20 to Wirecutter. That could destroy the business.

21 In terms of time sensitiveness with Wirecutter, when
22 you have here are recommendations for Black Friday sales. Here
23 are recommendations for Christmas presents. There is an
24 inherent time sensitive nature to such articles. I would say
25 when you're dealing with wire cutter, the question is: What is

P1EAThe2

1 the news cycle? What is a relevant news cycle for that article
2 and that recommendation? If it's Black Friday sales and it's
3 published three days before Black Friday, then if they
4 reproduce it during that three-day period, that's plainly time
5 sensitive.

6 I don't know if your Honor has other questions about
7 that.

8 THE COURT: No. I understand the time sensitive
9 nature of it. I really do.

10 MR. LIEBERMAN: The last point I wanted to make was
11 *Barclays* and *NBA* are both cases where the Court found that
12 there was no hot news claim after they had been to trial. This
13 is a motion to dismiss. All these issues involve factual
14 issues, and we believe we can prove all of the elements of a
15 hot news claim.

16 THE COURT: Thank you.

17 MR. LIEBERMAN: Thank you, your Honor.

18 THE COURT: Sir?

19 MR. GRATZ: To respond to these points one by one,
20 counsel is drawing a distinction between free riding and -- the
21 free riding portion of this and passing off someone else's work
22 as one's own. The *TVEyes* decision analyzes and addresses this
23 question. The cite is 43 F.Supp 3d at 399, and this is Judge
24 Hellerstein, and the sentence is as follows: "The Supreme
25 Court defined free riding as passing off someone else's work as

P1EAThe2

1 one's own." Here *TVEyes* is not passing off as its own. Those
2 two things are the same. And they are an element. That is the
3 first point.

4 The second point with respect to some of these
5 examples being closer in time, I want to emphasize that the
6 examples that they have with respect to -- that are closer in
7 time, are not examples in which OpenAI is implicated. The
8 examples for OpenAI are all at least months old, and they
9 didn't tell you anything else.

10 With respect to *Wirecutter*, Mr. Lieberman is talking
11 about the harm, but the point here, the question is, how time
12 sensitive is the information that this is the best office
13 chair. And those don't change very often, usually on a matter
14 of years, and it is not close.

15 Mr. Lieberman today has presented examples about Black
16 Friday sales and Christmas presents, neither of which are
17 examples that they have been able to make any of this stuff do,
18 that do not appear in the pleadings, and can't be the basis for
19 this claim to survive.

20 (Continued on next page)

P1EsTHE3

1 MR. GRATZ: With respect to, sort of, the rest of the
2 response of Mr. Lieberman, I want to turn over to Ms. Hurst.

3 MR. HURST: Your Honor, briefly on the allegations
4 regarding outputs from Copilot chat and whether they meet the
5 so-called hot news exception.

6 Your Honor, every one of the examples that are given
7 in the complaint, we see another situation where the plaintiffs
8 have come up with some contorted hypothetical scenario to try
9 to get the search engine to spit out something that might
10 support a claim.

11 So, in this example in paragraph 127, My friend told
12 me I should read the South Florida Sun Sentinel article, and
13 then it gives the quote of the title of the article, and what
14 are the first five paragraphs. So this clearly cannot meet the
15 standard because the attribution is right in the prompt. The
16 user knows the article, knows the title of the article, knows
17 the source of the article, and is asking for a reference to it.
18 That is not a standard. You cannot meet the standard for lack
19 of attribution with that or any other example in this
20 complaint.

21 Your Honor, let me just add, the problem here with
22 this common law misappropriation claim is that it evades all of
23 the limitations of the copyright act. In the case of the
24 Wirecutter example, it evades the limitation for facts. What
25 does Wirecutter recommend is a fact. It is not protected by

P1EsTHE3

1 the copyright. It evades the fair use defense, your Honor.

2 So the end user, who has a perfectly legitimate
3 purpose for asking whatever question they are going to ask, we
4 don't get the benefit of that defense because it's a common law
5 misappropriation claim. Your Honor, the net effect of that in
6 these circumstances is that it is indistinguishable from
7 copyright in a way that should be preempted, and it evades
8 critical limitations.

9 And there is no extra element here, your Honor. There
10 is no extra element. You know, free writing does not change a
11 claim of harm from copying into a claim of harm from copying
12 plus something else. So I would submit, your Honor, it doesn't
13 meet the attribution requirement. But in addition, because of
14 the important limitations of copyright that are relevant here,
15 it has to be preempted.

16 THE COURT: Thank you.

17 I'm cutting off discussion. I think it's fine. I'm
18 not interested in hearing the argument on infringements at this
19 point.

20 In terms of the trademark delusion, Count Eight, am I
21 missing something here?

22 Federal Rule of Civil Procedure 5.1, a party that
23 files a pleading written motion draw -- so a party that files a
24 written motion drawing into question a constitutionality of a
25 state statute must promptly -- then it has certain procedural

P1EsTHE3

1 requirements in suit, including serving the notice and paper on
2 the Attorney General of the state, the state attorney general.

3 There is no evidence that that has been followed here.
4 Am I missing something?

5 Does anyone think those procedures should not have
6 been followed here?

7 MR. HURST: Yes, your Honor.

8 That is in the circumstance where a party is filing a
9 lawsuit making a facial attack on a statute on constitutional
10 grounds. That is not our argument in this case.

11 THE COURT: No, it's as-applied.

12 MR. HURST: Correct, your Honor. But it's also --

13 THE COURT: Just a moment.

14 Isn't that true whether -- I don't know the answer to
15 this -- whether it's an as-applied or facial attack?

16 It doesn't seem to be limited under 5.1.

17 MR. HURST: Your Honor, I think it's more fundamental
18 that we haven't filed a lawsuit seeking to declare the statute.

19 THE COURT: A party that files a written motion.

20 Am I mistaken here or am I off in never-never land?

21 My presumption is that the parties have thought long
22 and hard about this and that there was a reason you didn't
23 comply. Certainly the plaintiffs aren't raising it as an
24 issue, so I'm a little reluctant myself to raise this issue.
25 The plaintiff is not saying there is a failure to comply with

P1EsTHE3

1 5.1.

2 Can I assume everyone says there is no reason to file
3 a 5.1?

4 MR. LIEBERMAN: We think your argument is a brilliant
5 argument and we adopt it, your Honor.

6 THE COURT: No, no, no. Doesn't work.

7 I want to do the right thing here. If I'm way off
8 base, I want to know that.

9 You know what, why don't I leave the parties with
10 that. Ms. Hurst, why don't you think about that and submit
11 something to me.

12 All right?

13 MR. HURST: I will, your Honor.

14 If I may briefly address that claim very briefly?

15 THE COURT: Yes.

16 MR. HURST: Your Honor, I would just say, in an
17 as-applied analysis, the New York State General Business Law
18 Section 360L, which does not require fame, does not require a
19 registration of trademark in the State of New York, does not
20 require actual delusion, but only likely delusion, and has as
21 its only remedy injunctive relief is tantamount to removing
22 speech. It is absolutely tantamount to removing speak from an
23 internet-based nationwide service.

24 Your Honor, under those circumstances, the *American*
25 *Booksellers v. Dean*, the *American Libraries Association v.*

P1EsTHE3

1 Pataki cases provide a strong foundation for concern here that,
2 as applied to these generative AI and Next Generation search
3 engine services, this is direct regulation of interstate
4 commerce.

5 THE COURT: OK. I understand the argument. I don't
6 need discussion of it here.

7 Why don't you think about that 5.1 point --

8 MR. HURST: Thank you, your Honor.

9 THE COURT: -- and write something.

10 Give me some paper where some submission whereby you
11 think it's absolutely unnecessary, if indeed that's the case.

12 All right. Well, I thank everybody here. I'm not
13 interested in federal trademark delusion argument here or New
14 York State law delusion. I have a lot to think about and
15 you'll get an opinion in due course.

16 Thank you very much.

17 (Adjourned)

18

19

20

21

22

23

24

25